

VOL. 166, NO. 1 THE AMERICAN NATURALIST JULY 2005

E-ARTICLE

A Null Model of Morphospace Occupation

Marcio R. Pie^{1,*} and Joshua S. Weitz^{2,†}

1. Department of Biology, Boston University, Boston, Massachusetts 02215;

2. Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544

Submitted September 21, 2004; Accepted March 15, 2005;

Electronically published April 21, 2005

ABSTRACT: Progress in understanding the relationship between lineage diversity, morphological diversity, and morphospace dynamics has been hampered by the lack of an appropriate null model of morphospace occupation. In this article, we introduce a simple class of models based on branching random walks (BRWs) for continuous traits. We show that many of the observed patterns of morphospace occupation might be simply a consequence of the dynamics of BRWs and therefore might not require special explanations. We also provide expected patterns of morphospace occupation according to a number of different conditions. In particular, we model BRWs on neutral landscapes and demonstrate that clumping in morphospace is possible even in the absence of adaptive landscapes with well-defined peaks and valleys. The quantitative definition of the BRW provides a means to analyze, both computationally and analytically, patterns of morphospace occupation according to different hypotheses.

Keywords: morphospace, random walk, evolution, diversification, null model, neutral theory.

Macroevolutionary studies over the past few decades have revealed several intriguing patterns. For instance, discordance between morphological and taxonomic diversity is common; that is, species richness of clades is often a poor predictor of morphological diversity (Foote 1993, 1996; Jernvall et al. 1996). Moreover, detailed studies have shown that morphological evolution tends to be accelerated during the early stages of the diversification of a group, increasing faster than lineage diversification but also decelerating sooner (Foote 1996; Rueber et al. 1999; Thomas

et al. 2000; Ciampaglio 2002). The search for the mechanisms underlying this deceleration in morphological diversification has generated considerable debate, particularly regarding its most celebrated example: the concentration of new animal phyla during the Cambrian explosion (Gould 1989; Conway Morris 1998).

There are two main types of explanations for this pattern (Erwin 1994; Valentine 1995). The empty ecospace hypothesis (Hutchinson 1959; Valentine 1980, 1995; Valentine and Walker 1986; Erwin et al. 1987; Valentine and Erwin 1987; Foote 1990), which traces back to Darwin (1859), suggests that evolutionary radiations begin in an unfilled “ecospace,” which becomes saturated as lineages diversify. In this light, the deceleration of morphological evolution could be interpreted to reflect a transition to an adaptive landscape marked by steeper peaks and deeper valleys (Raup 1966, 1967; Saunders and Swan 1984; Swan and Saunders 1987). Similarly, tighter morphological clusters would represent groups of species with more finely partitioned resources (Valentine 1969). Alternatively, the genomic hypothesis (Valentine 1986; Valentine and Erwin 1987; Arthur 1997) posits that developmental processes have become increasingly canalized (“entrenched”) over the history of life, therefore preventing further morphological diversification. These hypotheses have proved difficult to disentangle because no unique predictions about expected patterns of morphological diversification and morphospace occupation have been proposed (Erwin 1994).

In this article, we argue that progress in understanding the relationship between lineage diversity, morphological diversity, and morphospace dynamics has been hampered by a lack of an appropriate null model of morphospace occupation. Previous studies led to important advances in several fronts (Raup and Gould 1974; Slatkin 1981; Bookstein 1988; Foote 1996; Gavrillets 1999b; Ciampaglio et al. 2001). Formative work by Raup and Gould (1974) posited that seemingly nonrandom structure in the correlations of phenotypic traits could arise as a consequence of branching processes in morphospace, a conclusion also reached by Bookstein (1988), though the parameter regimes for producing heterogeneities and/or clustering were not explored. The importance of stochastic fluctuations in caus-

* Present address: Departamento de Zoologia, Universidade Federal do Paraná, 81531-990 Curitiba, Paraná, Brazil; e-mail: pie@ufpr.br. Order of authors was determined alphabetically.

† E-mail: jsweitz@princeton.edu.

ing differences in rates of lineage and morphospace diversification was also highlighted (Raup and Gould 1974), leaving open the question of the expected difference in diversification rates calculated from ensemble averages, that is, rates averaged over many realizations. The use of random walks was extended by Slatkin (1981) to analyze a diffusion-like model of morphospace dynamics in neutral and nonneutral environments. The problem with taking the diffusion limit of branching processes is that doing so eliminates fluctuations critical to generating realistic patterns of morphospace occupation; for example, the work of Young et al. (2001) notes the breakdown of the diffusion approximation in a different context.

Recently, a series of stochastic models have been developed with a shared aim of describing general patterns of morphospace dynamics of continuous and discrete character traits (Foote 1996; Gavrillets 1999b; Ciampaglio et al. 2001), with particular attention to the dynamics of adaptive radiations. The rapid increase of morphological disparity relative to lineage diversity is observed in these simulations, a consequence, in part, of the relatively slow growth in the variance of random walks relative to the often exponential increase in the number of surviving lineages. More detailed insights into whether heterogeneities in morphospace can generally be classified as indicative of selection, randomness, or constraints (or some combination thereof) are limited by the seeming multiplicity of factors and parameters that could generate related patterns. Analytical progress has been made by Gavrillets (1999b), who developed an elegant model of morphospace occupation that explained the commonly observed pattern of deceleration in clade diversification in terms of the geometric structure of morphospace without invoking major declines in the size of morphological transitions or taxonomic turnover rates. The reason for this effect is that as a clade expands into morphospace, it becomes less and less probable that a random morphological change will lead outside the volume of the morphospace already occupied by the clade. This model was extended to incorporate the presence of hard boundaries that limit the range of possible morphologies ("holey landscapes"), particularly in the case of discrete characters (Gavrillets 1999a). However, it would seem important to understand the dynamics of morphospace occupation in the case of continuous quantitative characters without such boundaries so that the basic impact of stochastic fluctuations in determining morphospace dynamics may be assessed.

To this end, we introduce a simple class of models based on branching random walks (BRWs). We show that many of the observed patterns of morphospace occupation might be simply a consequence of the dynamics of BRWs and therefore might not require special explanations. We also describe patterns of morphospace occupation according

to a number of different conditions. In particular, we model BRWs on neutral landscapes, and we demonstrate that clumping in morphospace is possible even in the absence of adaptive landscapes with well-defined peaks and valleys (Wright 1932). The quantitative definition of the BRW provides a means to analyze, both computationally and analytically, patterns of morphospace occupation according to different hypotheses. We begin by treating the simple case of a BRW without extinction and then consider how extinction, logistic diversification, developmental entrenchment, varying dimensionality, and mass extinctions might shape the characteristic features of morphospace dynamics.

Branching Random Walks

Random walks have been commonly used in modeling the evolution of quantitative traits in the case of evolution through drift or when the direction of selection changes randomly over time (Felsenstein 1985, 1988; Lynch 1990; Martins 1994). The models presented here are based on the branching random walk (BRW; Feller 1968, 1971; Dynkin 1991; Le Gall 1999; Young et al. 2001), a process where the number of walkers may change over time. In the context of this article, each walker describes one lineage represented by a point in a d -dimensional morphospace. A change in the position \vec{x}_i of the lineage in morphospace represents change in the mean value of phenotypic characters for that lineage. Branching and disappearance in the BRW represent speciation and extinction, respectively. In all BRWs that we examine, the change in the d -dimensional position in trait space of every extant lineage is governed by a continuous random walk of the form

$$d\vec{x}_i = \sigma(t)d\vec{W}_i(t), \quad (1)$$

where i is the index of the lineage, $\sigma(t)$ is the standard deviation of the incremental change in the mean phenotypic character, and $\vec{W}_i(t)$ is a stochastic Wiener process satisfying $dW_i(t')dW_j(t) = \delta_{ij}\delta(t' - t)dt$ (van Kampen 2001). It is important to note that BRWs should be applicable in the context of characters under directional selection, where the direction of selection changes randomly with time, as well as for the neutral evolution of characters, provided they are not influenced by epistatic interaction with characters under selection.

The study of branching processes (Kendall 1948; Harris 1963; Jagers 1975; Harvey et al. 1994) and random walks has been repeatedly applied to biological systems, and in our view, it seems natural to combine and apply these techniques to the problem of morphospace variation. Models of BRWs are parametrized by the speciation rate

b , the extinction rate m , and the diffusion rate $D \equiv \sigma^2/2$. The diffusion rate controls the change in mean phenotypic characters for extant lineages, while speciation and extinction determine the rate of change in the total number of lineages. In this simplified null model view, there are no peaks or valleys in the adaptive landscape and no character-dependent interactions between lineages; nonetheless, a number of realistic patterns emerges despite this seeming lack of complexity.

In this study, we assess the behavior of BRWs under a variety of conditions by calculating statistical features of the spatial distribution of lineages in concert with visual representations of morphospace occupation. The statistical analysis we implement includes measurements of mean pairwise distance between lineages $\mathcal{L}(t)$, total disparity $Z^2(t)$, and the cluster-size distribution $\hat{c}(\hat{r}, t)$. The mean pairwise distance,

$$\mathcal{L}(t) = \frac{2}{N(N-1)} \sum_{i,j>i}^N |\vec{x}_i(t) - \vec{x}_j(t)|, \quad (2)$$

and total disparity, $Z^2(t)$,

$$Z^2 = \frac{1}{N} \sum_{i=1}^N (\vec{x}_i - \vec{x}_0)^2, \quad (3)$$

capture the typical spread of phenotypic characters about the population mean, where

$$\vec{x}_0 = \frac{1}{N} \sum_{i=1}^N \vec{x}_i$$

is the average position of the lineages in morphospace. Note that Z^2 and the average pairwise distance squared are related by a simple algebraic relation, as derived in appendix A. The cluster-size distribution is a more detailed measure of how lineages are distributed in morphospace. For a given separation r , the cluster-size distribution is defined as the average size of clusters of lineages linked by distances less than r . When $r \rightarrow 0$, each lineage is a distinct cluster, whereas when $r \rightarrow \infty$, the entire set of lineages naturally groups into a single cluster. Discontinuous jumps in average cluster size at a given separation, as measured by the cluster-size distribution, provide a graphical means to detect multiple scales of organization indicative of nontrivial inhomogeneities (Plotkin et al. 2002). Detailed definitions and methodology to calculate these measures are provided in appendixes A and B. We explore the behavior of BRWs and demonstrate how they can be used as a suite of null models to assess how dynamical process relates to evolutionary outcome.

BRW without Extinction

The simplest model of a BRW is one with a fixed speciation rate, b , no extinction, $m = 0$, and a constant diffusion rate, D . This type of BRW is equivalent to imposing spatial diffusion atop a Yule process (Harris 1963). At $t = 0$, a single lineage is placed at the origin of a d -dimensional morphospace, representing a flat adaptive landscape for which change in any direction is equally likely. This lineage is the progenitor of all subsequent lineages, whose average number increases exponentially, $\langle N(t) \rangle = e^{bt}$. In simulations, as in rapid evolutionary diversification events, this exponential increase in total number of lineages is not expected to continue unbounded. To facilitate analysis of BRWs without extinction, we set a cutoff at $t = t_c$ so that $\langle N(t = t_c) \rangle = 500$.

The dynamics of lineages is initially characterized by a single lineage obeying a random walk in morphospace until the first speciation event. Subsequent lineages increase in number and cluster around the position of the first few branch points. At $t = t_c$, patterns in morphospace exhibited by BRWs without extinction show a single clump of points spread over a characteristic distance, $r_c \approx (dD \log N_c/b)^{1/2}$, without any significant scales of organization at smaller scales. Note that rescaling time, $t \rightarrow bt$, and space, $x \rightarrow (b/D)^{1/2}x$, collapses simulations with different parameters, b and D , onto the same universal pattern. Visualizations of simulations of BRWs with different values of b and D confirm this collapse, as does analysis of the clustering function, $\hat{c}(\hat{r}, t = t_c)$, which finds a threshold-like transition (see fig. 1 for details). A threshold-like transition implies that the lineages group essentially in a single clump reflecting their common phylogenetic history. Also, the average pairwise distance squared, $\langle \mathcal{L}^2(t) \rangle$, is nonlinear with a transition at $t \approx 1/b$. This result stands in contrast to the naive expectation that $\langle \mathcal{L}^2(t) \rangle$ is a linear function as it would be for random walks without branching. The nonlinearity arises because ensemble averages weight equally those branching processes with a different number of lineages, even though late diversification of a process implies stronger correlations among the traits of surviving lineages. A more detailed explanation for this phenomenon and supporting results from simulation can be found in appendix A. Note also that the mean pairwise distance, $\langle \mathcal{L}(t) \rangle$, rises more rapidly at early times than the expected number of lineages, $\langle N(t) \rangle$.

BRW with Speciation and Extinction

A BRW with extinction is a more realistic evolutionary model of the divergence of lineages than that of the previous section ("BRW without Extinction"). Because of the stochastic nature of the process, a BRW with fixed speciation rate b and extinction rate m has a finite probability

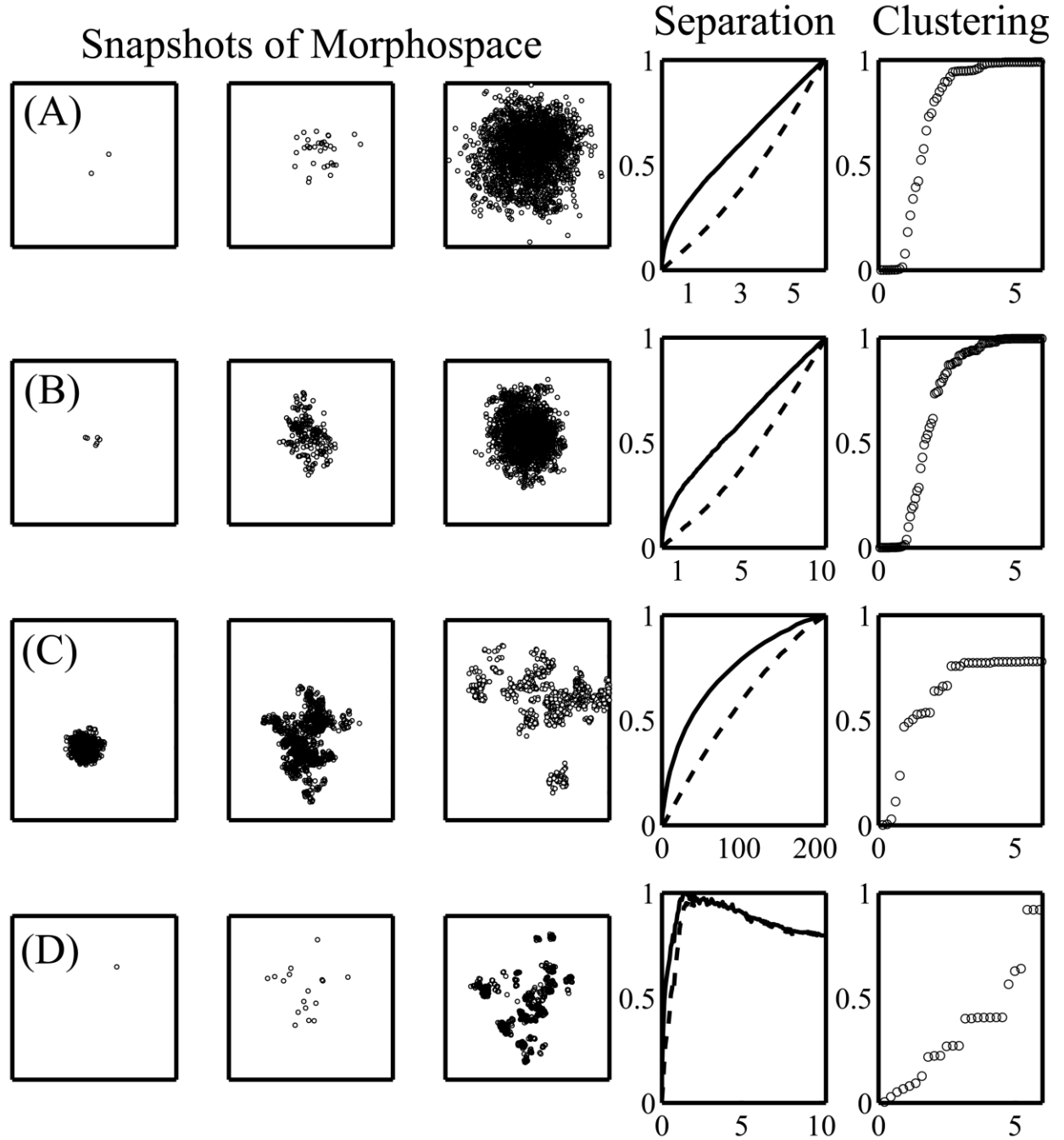


Figure 1: Depiction of the behavior of BRWs under different conditions: (A) BRW without extinction, (B) BRW with speciation and extinction, (C) BRW with logistic diversification, (D) BRW with developmental entrenchment. A series of snapshots of a typical BRW are depicted in the first three figures corresponding to each type of process. The mean pairwise distance $\mathcal{L}(t)$ (solid lines) and mean pairwise distance squared $\mathcal{L}^2(t)$ (dashed lines) averaged over $3 \times 10^2 - 10^3$ realizations in the column labeled “Separation.” Finally, a single calculation of the cluster size distribution of the morphospace occupation is depicted in the column labeled “Clustering.” In all cases, $\mathcal{L}(t)$ and $\mathcal{L}^2(t)$ increase with time during early diversification, but in the case of developmental entrenchment, these measures actually decrease. Normalization is with respect to the maximum value of separation and disparity, and the time axis is dimensionless, $t \rightarrow bt$. Notice also that the clustering of the BRW both with logistic diversification and with developmental entrenchment differs qualitatively from the simple BRWs with $b > m > 0$ and with $b > m = 0$.

$p_s = 1 - m/b$ of persisting indefinitely if $b > m$ (Kendall 1948). We therefore consider only those processes for which $b > m$, since otherwise the diversification process is ensured to end in complete extinction. For the supercritical case, $b > m$, the average number of lineages increases exponentially, $\langle N(t) \rangle = e^{(b-m)t}$.

Analysis and visualizations of BRWs with $b > m$ find patterns similar to those with $b > 0$ and $m = 0$, that is, a single clumped pattern in morphospace. This is unsurprising given that both are in the supercritical regime, $b > m$. The elimination of lineages has many interesting consequences for reconstructing phylogenies (Harvey et al. 1994; Nee et al. 1994), whereas here we are concerned with the spatial distribution of extant lineages in morphospace instead of the temporal distribution of branch points, though the two are certainly related (Edwards 1970). Rescaling the domain of morphospace distributions based on the upper bound of total disparity, $Z^2(t)$, suggests that the presence of extinction merely retards the inevitable expansion of lineages outward into morphospace. It would appear that simple BRWs with $b > m$ can account for the relatively rapid increase in separation as compared with growth in lineages.

BRW with Logistic Diversification

A longstanding debate in the literature concerns whether lineage diversification is exponential or saturated (Benton 1997; Sepkoski and Miller 1998). The simplest model of lineage saturation in a BRW is to describe the dynamics of lineages via logistic growth (Walker and Valentine 1984),

$$\frac{dN}{dt} = bN \left(1 - \frac{N}{K} \right), \quad (4)$$

where $m(N) = bN/K$ and K is the carrying capacity of lineages in morphospace. The pattern of morphospace occupation in steady state can be markedly different from that depicted in figure 1. At the fluctuating steady state, the distribution of lineages in morphospace is typified by well-separated clusters. For example, consider the results of a typical simulation of a BRW with logistic growth, for which $K = 1,000$. The time development of morphospace occupation is depicted in figure 2. Initially, the lineage spreads out like a BRW without extinction (since $m(N) \ll b$ when $N \ll K$). But once the number of lineages approaches saturation, the extinction rate becomes approximately equal to the speciation rate; heterogeneities in spatial distributions are reinforced and introduced because of stochastic fluctuations in the BRW.

These fluctuations notwithstanding, how does the clustering of lineages arise in the absence of selective pressures

or niche-based interactions? The answer lies in the interplay between branching and diffusion in the BRW; although diffusion smooths out heterogeneities, the branching process intensifies them. The extinction of lineages eliminates intermediates in morphospace while birth reinforces local clumping (because birth is by nature a multiplicative process centered around the mother). The mechanism for clustering is similar to that outlined by Young et al. (2001), though with drastically different initial conditions.

The reason for clustering is as follows. Upon reaching steady state, imagine that the $N \approx K$ lineages are distributed homogeneously over a morphospace volume V in d dimensions. The length scale over which a mother and daughter lineage separates before a subsequent branching is $l_r = (dD/b)^{1/2}$. If mother and daughter lineages are unable to diffuse across typical separation distance $l_p = (V/N)^{1/d}$, then clusters develop. Therefore, clustering becomes increasingly likely as $l_p > l_r$. At $t = 476$ of figure 2, $l_p/l_r \approx 1.5$, confirming that we are in the clustering regime, whereas at $t = 6$, $l_p/l_r \approx 0.35$ in accord with the lack of clustering.

BRW with Developmental Entrenchment

Another means of making a BRW more realistic is to modify the condition that the diffusion rate remains constant. A BRW with developmental entrenchment assumes that the diffusion rate in morphospace decreases monotonically with time; that is, $(dD(t)/dt) < 0$. This decrease in diffusion simulates inability of organisms to modify character traits after initial diversification. Unsurprisingly, the morphospace occupation of such a BRW possesses the generic feature of clustering after an initial period of rapid expansion and diversification.

A quantitative analysis of this clustering was conducted for the specific case of

$$D(t) = D_0 \times e^{-[(b-m)t]/C}, \quad (5)$$

where D_0 is arbitrary, $b = 0.0069$, $m = b/2$, and $C = 1.4$. Using these parameters, the mean cluster size as a function of dimensionless radius was analyzed for a set of 10^3 ensembles. The results are presented in figure 3, where the mean cluster size \hat{c} is shown as a function of normalized lineage separation \hat{r} . The curve rises rapidly, implying the presence of many small clusters, and then increases slowly as $\hat{r} > 1$. The clustering curve is contrasted with a curve derived from homogeneously distributed data possessing only a single scale of organization, that is, that set by the density of lineages. In the case of a BRW with developmental entrenchment, there is no single scale defining the

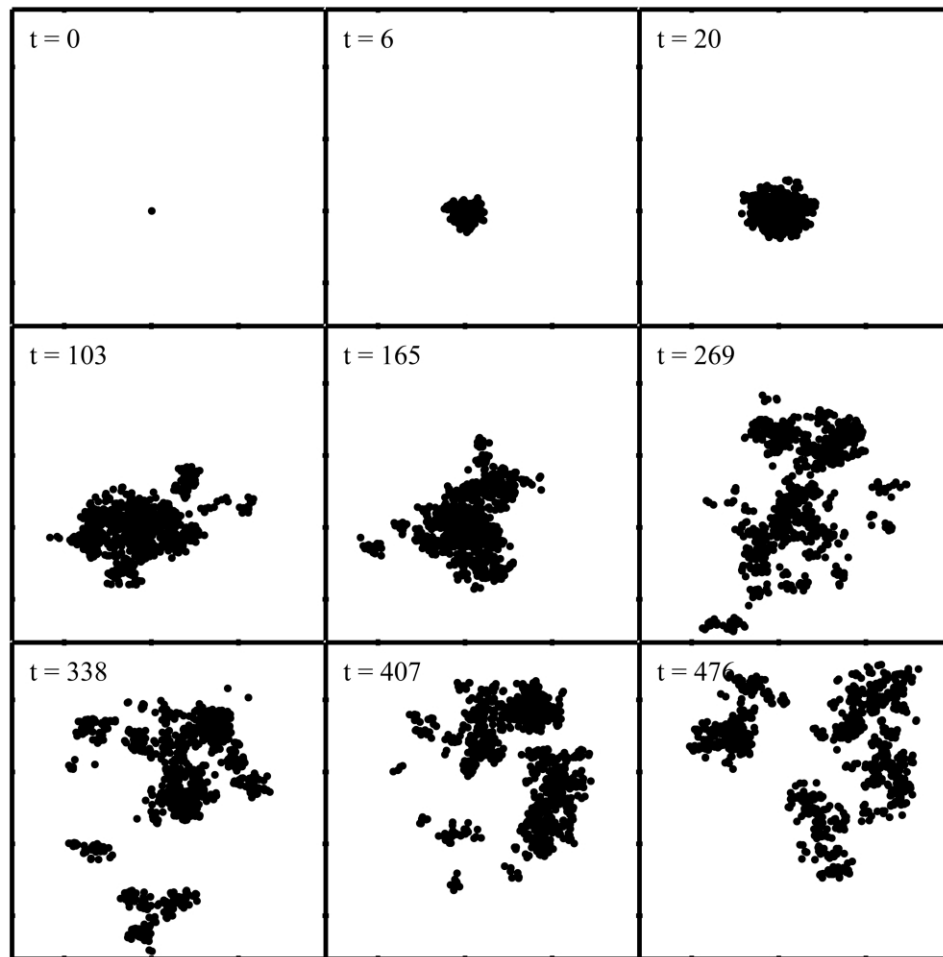


Figure 2: Series of snapshots of a typical BRW with logistic growth: $b = 0.069$ and $K = 1,000$. The clustering of lineages in morphospace occurs after a dynamic equilibrium in the number of lineages has been reached. In all cases, the time is labeled in dimensionless units, and the X- and Y-axes are the same for each plot.

points in morphospace; multiple clusters are a generic property of the long-term dynamics.

The explanation for the clustering is apparent when one imagines the morphospace dynamics for a series of lineages in the regime $D(t) \rightarrow 0$. In this slow-diffusion limit, individual points do not mix, and hence fluctuations will lead to the disappearance of some regions and the reinforcement of others; that is, $l_p \gg l_r$, as explained in “BRW with Logistic Diversification.” The effect is further intensified in the case of logistic growth. Note that in the long-time limit, as $D(t) \rightarrow 0$, the reasons for clustering are nearly the same as those espoused for clustering in the case of spatially diffusing plankton (Young et al. 2001), though we do not restrict ourselves to the special case $b = m$.

BRW in Higher Dimensions

Although descriptions of BRWs so far focused on two dimensions, these methods apply equally well to higher dimensions. A BRW in higher dimensions with $b > 0$ and $m = 0$ will show a nonlinear increase in mean pairwise distance squared, though with an increased prefactor, $d^{1/2}$. The same is true for $b > m > 0$, that is, a BRW with speciation and extinction. The expected number of lineages is not affected by the dimensionality of the system, though this does not preclude hypotheses where the dimensionality of the morphospace is coupled to the saturating number of lineages.

A possible distinction between $d \leq 2$ and $d > 2$ concerns whether the presence of clustering satisfies the definition

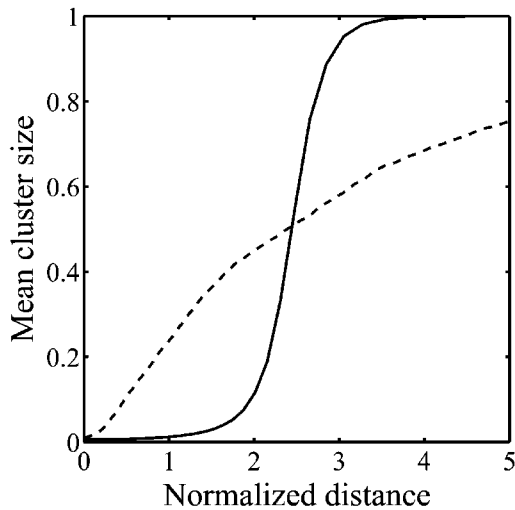


Figure 3: Mean cluster size $\hat{c}(\hat{r})$ as a function of normalized separation \hat{r} for the case of randomly distributed data (solid line) and ensemble-averaged BRWs with developmental entrenchment (dashed line). The BRWs all have $b = 0.0069$ and $m = b/2$, and the diffusion rate decays over the dimensionless timescale $t = 0.35$. Notice that the slow increase of the cluster size function for BRWs with developmental entrenchment reflects the many small and intermediate clusters that develop as a result of the slowing down of diffusion in morphospace.

of a phase transition, in the sense that spatial fluctuations increase in an unbounded fashion with increasing time. A recent analysis of a spatial lattice model with fluctuation-induced clustering (Houchmandzadeh 2002) where $b = m$ finds that fluctuations can become arbitrarily large in dimensions $d > 2$. The same article (Houchmandzadeh 2002) also concludes that “clustering” is precluded in the strict, mathematical sense of divergent fluctuations but permits the practical definition of $l_p > l_r$ described earlier. The higher dimensional systems can be thought of as less constrained topologically and less likely to demonstrate clustering. In the case of a BRW with developmental entrenchment, the strict mathematical definition for clustering is satisfied, even in $d > 2$. When the diffusion rate diminishes toward 0, the higher dimensional system behaves like a system of disconnected spatial patches where clustering is expected. The case of logistic diversification seems a natural subject for future work. Much of the current work on fluctuation-induced clustering (Houchmandzadeh 2002; Fuentes et al. 2003; Shnerb 2004) could easily be applied and extended to the initial conditions and evolutionary models considered in the study of morphospace divergence from a single lineage.

BRW with Mass Extinctions

In the context of BRWs, a mass extinction event is considered to be the simultaneous extinction of multiple lin-

eages at a given instant in time. Consider a group of N lineages at time t with continuous trait values \hat{x}_i . The simplest type of extinction event is one in which a random fraction, $0 < p < 1$, of lineages is killed irrespective of their phenotypic traits, leaving only $N' \approx (1 - p)N$ remaining. Such an extinction inexorably alters the occupation of morphospace, but does it do so in a statistically significant manner? For example, how does the average pairwise distance squared, $\langle \mathcal{L}^2 \rangle$ between lineages, differ before and after such an event? Random subsampling of lineages will not alter the underlying spatial distribution of morphospace occupation and its derived spatial statistics (ignoring corrections on the order of $1/N$). Therefore, the expected value of \mathcal{L}^2 should be approximately equal before and after an extinction event, despite the sparser coverage of morphospace.

We may then imagine running forward in time the same set of dynamics that generated the preextinction pattern in morphospace with the remaining lineages. What happens once the number of extant lineages has returned to the preextinction level of N ? If $\langle \mathcal{L}^2(t) \rangle$ increases with time for the BRW (as it does for all BRWs considered except for those with developmental entrenchment), then the mean pairwise distance squared should increase. Thus, the expected value of \mathcal{L}^2 will be greater at $t + \tau$ than at t , where τ is the time necessary for the number of lineages to increase from N' to N . The increase of \mathcal{L}^2 occurs despite the fact that the number of lineages drops precipitously at a given time or sequence of times. Results from a typical simulation with a sequence of three extinction events with $p = 0.9$ are shown in figure 4. The passage of time, not the number of mass extinction events, is the dominant driver of the increase in morphological disparity. This by no means catalogs the effect of mass extinctions on all aspects of morphospace occupation. Rather, it provides an expectation of how morphospace structure is affected by nonselective mass extinctions.

Discussion

Branching random walks are a class of models that can be used to provide null expectations of patterns of morphospace occupation under different conditions. The impetus for this exploration of BRWs was a seminal article by Raup and Gould (1974) that explored a model similar to our own, namely a branching process with fixed extinction and diversification rates, and exclusively punctuated evolution where only one of the daughter lineages was allowed to evolve through discrete steps in morphological space. Interestingly, their findings include an irregular occupation of morphospace reflecting the diffusion of character traits that share a common phylogenetic history. The nonhomogeneous occupation arises because re-

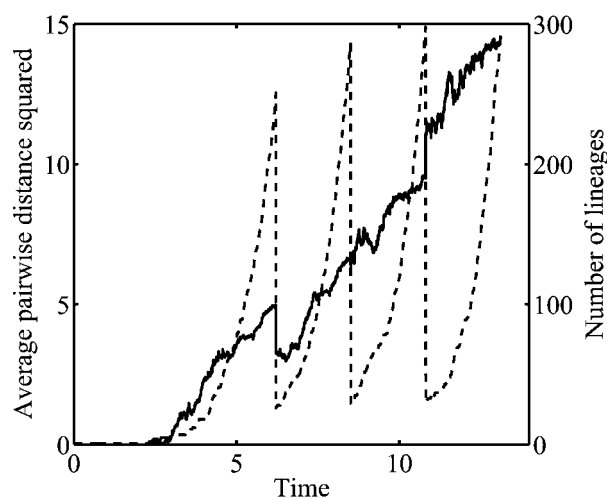


Figure 4: Average pairwise distance square, $\mathcal{L}^2(t)$ (solid line), and number of lineages, $N(t)$ (dashed line), as a function of dimensionless time for a simulation of a BRW without extinction but with a sequence of three mass extinction events. The mass extinction events for which a fraction $p = 0.9$ of lineages is eliminated occur at dimensionless times 6.2, 8.5, and 10.8. Notice that $\mathcal{L}^2(t)$ continues to increase after each mass extinction event.

lated lineages undergo correlated random walks; although the resulting character traits of any given simulation are random, their relatedness is inevitable. However, our results indicate that the formation of discrete clusters in morphospace is not characteristic of all BRWs, being present only in cases of logistic lineage diversification and in the presence of developmental entrenchment.

Foote (1990) extended the Raup-Gould approach to a wider range of conditions. His models also used exponential lineage growth and discrete character transitions but focused on morphological diversity on a single dimension. Several evolutionary scenarios were explored, including logistic lineage diversification, developmental entrenchment, and variation in lineage turnover. Interestingly, our results differ from those obtained by Foote in two main issues. First, in the case of a simple BRW with exponential diversification, we observed that average pairwise distance squared increases nonlinearly over time. Second, developmental entrenchment in our model caused an increase, followed by a waning and finally a constant level of morphologically diversity, a pattern surprisingly similar to the dynamics of the Cambrian explosion described by Gould (1989). Conversely, Foote's results indicated an asymptotic increase in morphological diversity, leading him to conclude that it was not possible to decide, on the basis of taxonomic diversity and morphological disparity alone, the reasons for the temporally and hierarchically heterogeneous deployment of disparity. The reason for this discrepancy is simple. Although both models used a monotonic decrease in the rate of morphological

diffusion to simulate developmental entrenchment, even in the smallest morphological steps in Foote's model, the typical length scale over which a lineage diffused in morphospace before a speciation/death event was still large in relation to the average distance between lineages. In other words, morphospace diffusion dominated lineage diversification as the leading factor in determining morphospace structure, as explained above. Had Foote chosen another monotonically decreasing function and allowed the simulation to run longer, he would probably find results similar to our model. This difference highlights the utility of analyzing the structure of morphospace in addition to variation in overall disparity over time.

Surprisingly, several aspects of the dynamics of morphospace occupation might not necessarily require special explanations given a BRW with extinction. This conclusion was anticipated by Foote (1990), who suggested that a progressive increase in the apparent nonrandomness of some aspects of morphospace occupation may be an expected consequence of diversification, with or without overriding ecological changes. One example is the (apparent) preferential elimination of lineages at the extremes of a morphospace (Foote 1993), a pattern often interpreted as evidence for higher susceptibility to extinction of specialized lineages (as indicated by their marginal position in morphospace). As shown in this article (see fig. 2 for an example), the relative concentration of lineages in certain areas of morphospace is not necessarily the product of an underlying adaptive landscape but rather an expected outcome of some simple classes of BRWs. Clustering can

arise as a consequence of extinction occurring everywhere in morphospace but speciation always occurring next to an extant lineage. Therefore, in order to demonstrate that the positioning of lineages in morphospace is caused by selective pressures, one needs to demonstrate that the observed patterns are significantly different from what would be expected by chance alone. Demonstrating the influence of adaptive landscapes on patterns of morphospace occupation might therefore be very difficult to demonstrate empirically, a conclusion also reached by Gavrillets (1999b). It is worth noting that repeated simulations of BRWs that display “clumping” should not preferentially cluster in the same portion of morphospace; that is, the average trait value of clusters is itself a random variable, which suggests a future avenue for comparison between BRW models and data.

Any comparison between theory and paleontological data is made more difficult by the often incomplete and noisy fossil record. These properties hinder detailed quantitative analyses that are often necessary to discriminate alternative macroevolutionary hypotheses. An important new area of research is the combination of neontological data with a time frame provided by molecular phylogenies. For instance, Harmon et al. (2003) analyzed four large lizard radiations to test whether morphological diversity was most strongly partitioned within rather than between subclades, a result that would be consistent with a disproportionately faster rate of morphological diversification early in the history of those clades. These results were then compared with a null expectation based on the simulation of character evolution on the inferred phylogenies using a Brownian model. Interestingly, different lizard clades showed varying levels of subclade partitioning of morphological diversity, a pattern that seems to be regulated by the dynamics of lineage diversification. The formalism of BRWs presented in this article can provide a firm framework for future work where such phenomena can be systematically explored.

Branching random walks might also be useful in the analysis of morphospace occupation during adaptive radiations. Erwin (1994; see also Ciampaglio 2002) suggested that measuring morphological disparity before and after mass extinctions could provide a test of the ecospace and genomic hypotheses. According to Erwin, extensive morphological innovation occurs after mass extinctions because of the availability of ecospace. On the other hand, morphological diversification is limited under the genomic hypothesis. Our model suggests that these predictions depend on the branching process. First, if the species that survive a mass extinction are a random sample of the preexisting species, the statistical structure of the morphospace remains unchanged though it is sparser (see fig. 4 for an example), confirming previous verbal arguments

(Foote 1993). In fact, detailed analysis of crinoid and blastozoan morphological diversity before and after three mass extinctions (end-Ordovician, end-Devonian, and end-Permian) suggests that morphospace structure is indeed conserved across mass extinctions (Ciampaglio 2002). This says nothing of the postextinction dynamics, which may well be limited because of entrenchment or may simply refill the sparse morphospace made available by the extinction.

Several authors have suggested that rapid phenotypic diversification can be caused by an increase in morphospace dimensionality (Erwin 1994; Kauffman 2000). For instance, Erwin (1994) suggested that the Cambrian explosion involved a process where morphospace was expanding in dimensionality during the radiation, and this expansion promoted further expansion through a “positive feedback” mechanism. Although the dynamics leading to the increase or decrease of dimensionality of morphospace cannot be directly inferred from the model presented here, our results confirm a slightly modified version of the previous verbal arguments. The rate of diffusion in morphospace increases with dimensionality in proportion to $d^{1/2}$, suggesting that the largest increases in the rate of morphospace occupation occur when additional dimensions are added to less complex organisms. Interestingly, this observation indicates a previously unrecognized trade-off: although the rate of morphospace diffusion increases with dimensionality, the rate of adaptive evolution decreases with it. High dimensionality has been recognized as hindering adaptive evolution, since Fisher (1930) suggested that the chance that a mutation of a given size will be favorable declines with the complexity of an organism, where complexity is defined as the number of morphospace dimensions. This issue has been recently addressed by Orr (2000), who showed that the magnitude of the cost of complexity is even higher than Fisher’s analysis had suggested. By taking into account the fact that favorable mutations must escape stochastic loss when rare (Kimura 1983), Orr used a population genetic model to show that the rate of adaptation of an organism declines as fast as d^{-1} . Therefore, the trade-off described here would predict that maximum long-term evolvability of a clade would occur at intermediate dimensionality. Moreover, mechanisms to decrease morphospace dimensionality, such as canalization, epistasis, and modularity (Wagner and Altenberg 1996), may be not only necessary but also fundamental components to allow adaptive evolution by acting as modulators of evolvability.

Acknowledgments

We thank J. Dushoff, M. Foote, C. J. Schneider, and an anonymous reviewer for helpful comments on the man-

uscript. M.R.P. was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior/Brazilian Ministry of Education, Brasília, Brazil. This manuscript is based on work supported by the National Science Foundation under a grant awarded in 2003.

APPENDIX A

Spatial Spread of a Branching Random Walk

Studies of morphological character evolution often classify the degree of spread/difference of a collection of lineages based on the average pairwise distance between phenotypic traits (e.g., Ciampaglio 2002). In this appendix, we calculate a number of basic measures related to the spatial spread of branching random walks, including but not limited to average pairwise distance.

Consider the divergence of a population of lineages that share a single, common ancestor. Intuitively, it would seem that at early times the average pairwise distance between lineages would be small, reflective of their common origins. As time increases, so too does the spread of phenotypic traits in a neutral morphospace. Is it possible to move beyond such qualitative claims? Fortunately, the answer is yes, though exact results are, at times, difficult to derive. In so doing, we focus on three related measures of spread: the total disparity, $Z^2(t)$; the pairwise distance between lineages, $\mathcal{L}(t)$; and the pairwise distance squared between lineages, $\mathcal{L}^2(t)$.

Population Measures of Spread

The spread of lineages in morphospace may be viewed as a stochastic, individual-based approach to the problem of population spread (Skellam 1951; van den Bosch et al. 1990). The analysis of the spatial spread of populations is typically analyzed by examining changes in density, where $\rho(\vec{x}, t)dV$ is the expected number of points within a volume dV centered around the position \vec{x} at time t . The equation describing the density dynamics is

$$\frac{d\rho}{dt} = D\nabla^2\rho + (b - m)\rho, \quad (\text{A1})$$

where D is the diffusion constant. Given an initial density of points at the origin, a density front grows outward in a spreading process with asymptotic velocity $v_s = [D(b - m)]^{1/2}$. However, as pointed out by Young et al. (2001), the continuum approximation is of limited use for analyzing spatial patterns when $b = m$. Additionally, the continuum approach is not even applicable in the case of a small number of lineages.

We therefore develop a series of measures that accounts

for the divergence of individual lineages. When $b > m$, a spreading process persists with probability $p_s = 1 - m/b$ (Kendall 1948). In such a case, what is the typical region over which the surviving particles are distributed? We define the total disparity, $Z^2(t)$, to be the measure of this spread. For a given realization,

$$Z^2 = \frac{1}{N} \sum_{i=1}^N (\vec{x}_i - \vec{x}_0)^2, \quad (\text{A2})$$

where

$$\vec{x}_0 = \frac{1}{N} \sum_{i=1}^N \vec{x}_i$$

is the average position of the lineages in morphospace. After some algebraic manipulation, the equation for Z^2 may be written as

$$Z^2 = \frac{1}{N^2} \sum_i^{N-1} \sum_{j=i+1}^N (\vec{x}_i - \vec{x}_j)^2. \quad (\text{A3})$$

Hence, the total disparity of a BRW may be recast in terms of the average pairwise distance squared between lineages. Analytical methods to solve for this expression are derived in the following.

Mean Pairwise Distance and Pairwise Distance Squared

Consider a set of N lineages with d -dimensional positions, $\{\vec{x}_i\}$. The pairwise distance between lineages is

$$\mathcal{L}(t) = \frac{2}{N(N-1)} \sum_{i,j>i}^N |\vec{x}_i(t) - \vec{x}_j(t)|, \quad (\text{A4})$$

and the pairwise distance squared is

$$\mathcal{L}^2(t) = \frac{2}{N(N-1)} \sum_{i,j>i}^N (\vec{x}_i(t) - \vec{x}_j(t))^2. \quad (\text{A5})$$

The relevant calculation for studies of morphospace occupation is to predict the ensemble-average quantities, $\langle \mathcal{L}(t) \rangle$ and $\langle \mathcal{L}^2(t) \rangle$. The difficulty in taking such an ensemble average is that $N(t)$ is itself a fluctuating quantity. For any given realization, it is worth noting the useful relation,

$$\mathcal{L}^2 = \frac{2N}{N-1} Z^2, \quad (\text{A6})$$

since it is computationally simple to calculate Z^2 .

It is important to note that $\langle \mathcal{L}^2(t) \rangle$ is not proportional to t for a simple BRW as would be the case for a random walk with a fixed number of lineages. The basic reason

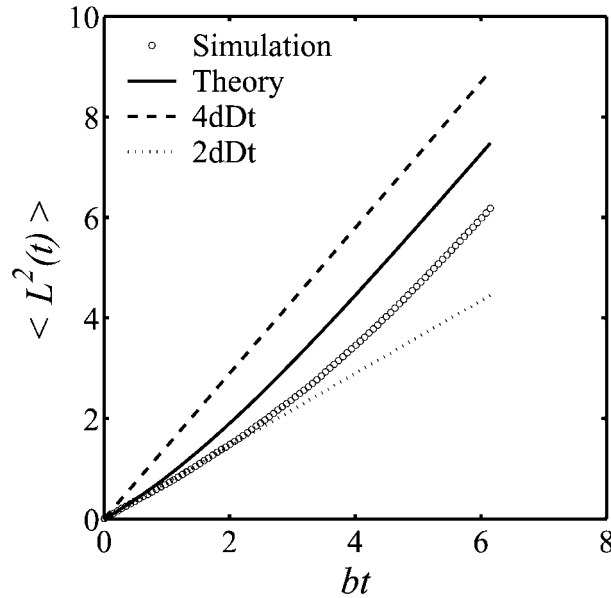


Figure A1: Average pairwise distance squared, $\langle \mathcal{L}^2(t) \rangle$, as a function of the dimensionless time bt is compared with theory and limiting curves. The solid line is the prediction of equation (A8). Lower dotted line is that of the average displacement squared of a single random walker. Upper dashed line is that of the average squared separation of two random walkers who begin together at the origin and is equal to the expected average pairwise distance squared for random walks without branching.

for the nonlinearity is the presence of correlations among lineages. An arbitrary particle pair $\tilde{x}_i(t)$ and $\tilde{x}_j(t)$ share a common spatial and phylogenetic history until $t = \tau$, where τ is the latest time at which i and j share a common ancestor. When $t > \tau$, the spatial component of the branching random walk is decoupled for the pair. Therefore, the average distance between such a pair is proportional to $[dD(t - \tau)]^{1/2}$, where d is the dimension and D is a morphospace diffusion constant. Likewise, the average distance squared between such a pair is exactly $4dD(t - \tau)$, that is, twice the variance of a single random walk.

In order to make progress analytically for BRWs, we consider the case of $b > 0$ and $m = 0$, that is, the branching random walk without extinction. A pair of lineages may be viewed as two lines in trait space with a single branch point (Hansen and Martins 1996). An expression for the average pairwise distance squared is

$$\langle \mathcal{L}^2(t) \rangle = \frac{\sum_{n=2}^{\infty} P_n(t)}{1 - P_1(t)} \int_0^t d\tau 4dD(t - \tau) Q(\tau; n(t), b), \quad (\text{A7})$$

where $P_n(t)$ is the probability of finding n lineages at time t and $Q(\tau; n(t), b)$ is the probability of an arbitrary pair of lineages separating at time $\tau < t$ given that the branching process resulted in n lineages at time t with branching rate b . An expression for $P_n(t)$ is known (Harris 1963), while the authors are unaware of an analytical expression for

$Q(\tau; n(t), b)$. We attempt to approximate $\langle \mathcal{L}^2(t) \rangle$ by assuming that every pair of lineages acts independently. In the absence of correlations, the probability a branch point occurred within $d\tau$ of $t = \tau$ is $dtbe^{-b\tau}/(1 - e^{-bt})$, and so we write

$$\begin{aligned} \langle \mathcal{L}^2(t) \rangle &\approx \int_0^t d\tau \frac{be^{-b\tau}}{1 - e^{-bt}} 4dD(t - \tau), \\ &= \frac{4dD}{b} \left(\frac{bt}{1 - e^{-bt}} - 1 \right). \end{aligned} \quad (\text{A8})$$

This approximation exceeds that of the ensemble average $\langle \mathcal{L}^2(t) \rangle$ from simulations and agrees with a biased average that weights simulations with N lineages by $N(N - 1)$. The actual ensemble average weights all processes equally regardless of N . Comparison of equation (A8) with numerical analysis is presented in figure A1.

There are certain features of the theoretical approximation and simulation results in figure A1 that are worth noting. The first is that it is an asymptotically exact result for $bt \ll 1$, where we can ignore the probability that more than one branch point will occur. The next interesting point is that when $bt \ll 1$, $\langle \mathcal{L}^2(t) \rangle = 2dDt$, that is, the same as for the spread of a single walker diffusing from the origin. However, when $bt \gg 1$, $\langle \mathcal{L}^2(t) \rangle \rightarrow 4dDt$, that is, the same as for the spread of two walkers diffusing from the

origin. Despite the fact that this theory is not yet exact, it suggests that the average pairwise distance squared should be nonlinear with a transition in slopes near $t \approx 1/b$. Likewise, the average pairwise distance should not be exactly proportional to $t^{1/2}$.

APPENDIX B

Clustering Metrics for Branching Random Walks

Statistical physics has long used the method of continuum percolation to cluster points (Stauffer and Aharony 1992). The notion of percolation is that there exists a connected chain of points, where the definition of “connected” depends on the phenomena under investigation. Methodologically, the idea is to connect all points for which there is a connected chain of points that have interparticle distances less than r . In more detail, the method is the following. First, calculate the interparticle distances between all points in the set $\{\vec{X}\}$ and denote these distances by d_{ij} . Second, for every set of points i and j , evaluate whether they are in the same cluster, $\chi_{ij} = \{0, 1\}$. Third, points i and j are in the same cluster if $d_{ij} < r$ or if there exists a set of points k_1, k_2, \dots, k_n such that $d_{ik_1} < r, d_{k_1k_2} < r, \dots, d_{k_{n-1}k_n} < r$.

For a given set of points $\{\vec{X}\}$ and a given scale r , the number of clusters will decrease from $n_c = N$ when $r \rightarrow 0$ to $n_c = 1$ when $r \rightarrow \infty$. For a set of points with density ρ , the number of clusters will undergo a transition at

$$r \approx \frac{1}{2\sqrt{\rho}}. \quad (\text{B1})$$

Likewise, the mean cluster size will go from $c(r) = 1$ to $c(r) = N$ as r goes from 0 to ∞ .

The mean cluster size has been used in forest ecology studies to analyze clustering of tropical trees (Plotkin et al. 2002). Nontrivial spatial correlations generate a staircase pattern in the mean cluster size $c(r)$ as opposed to a percolation-like threshold typically observed in random data. The plateaus themselves denote scales of nontrivial organization aside from that set by $1/\rho^{1/2}$.

This clustering technique provides a standard means to compare data sets with differing number of lineages N , which is especially useful in the case of branching random walks. If the data has been clustered into M clusters, each with c_i number of points in each cluster, then the mean cluster size is

$$\langle c \rangle = \frac{1}{N} \sum_i^M c_i^2. \quad (\text{B2})$$

The mean cluster size is normalized, $\hat{c} = \langle c \rangle / N$, to extend

from 0 to 1. The normalized distance is $\hat{r} = r/r_{NN}$, where r_{NN} is the average nearest neighbor distance defined to be

$$r_{NN} = \frac{1}{2} \sqrt{\frac{V}{N}}, \quad (\text{B3})$$

where V is the morphometric volume covered. The morphometric volume, V , is defined in all cases to be the area/volume/hypervolume of the smallest circle/sphere/hypersphere that circumscribes all extant points in morphospace. All plots of clustering presented in this article are displayed in terms of \hat{c} versus \hat{r} .

Literature Cited

- Arthur, W. 1997. The origin of animal body plans. Cambridge University Press, Cambridge.
- Benton, M. 1997. Models for the diversification of life. *Trends in Ecology & Evolution* 12:490–495.
- Bookstein, F. 1988. Random walks and the biometrics of morphological characters. *Evolutionary Biology* 23:369–398.
- Ciampaglio, C. 2002. Determining the role that ecological and developmental constraints play in controlling disparity: examples from the crinoid and blastozoan fossil record. *Evolution and Development* 4:170–188.
- Ciampaglio, C., M. Kemp, and D. McShea. 2001. Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity. *Paleobiology* 27:695–715.
- Conway Morris, S. 1998. Crucible of creation: the burgess shale and the rise of animals. Oxford University Press, Oxford.
- Darwin, C. 1859. On the origin of species. J. Murray, London.
- Dynkin, E. 1991. Branching particle-systems and superprocesses. *Annals of Probability* 19:1157–1194.
- Edwards, A. 1970. Estimation of the branch points of a branching diffusion process. *Journal of the Royal Statistical Society B* 32:155–174.
- Erwin, D. 1994. Early introduction of major morphological innovations. *Acta Palaeontologica Polonica* 38:281–294.
- Erwin, D., J. Valentine, and J. Sepkoski. 1987. A comparative study of diversification events: the early Paleozoic versus the Mesozoic. *Evolution* 41:1177–1186.
- Feller, W. 1968. An introduction to probability theory and its applications. Vol. 1. 3rd ed. Wiley, New York.
- . 1971. An introduction to probability theory and its applications. Vol. 2. 2nd ed. Wiley, New York.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* 125:1–15.
- . 1988. Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics* 19:445–471.
- Fisher, R. A. 1930. The genetical theory of natural selection. Oxford University Press, Oxford.
- Foote, M. 1990. Nearest-neighbor analysis of trilobite morphospace. *Systematic Zoology* 39:371–382.
- . 1993. Discordance and concordance between morphological and taxonomic diversity. *Paleobiology* 19:185–204.
- . 1996. Models of morphological diversification. Pages 62–86

- in D. Jablonski, D. Erwin, and J. Lipps, eds. *Evolutionary paleobiology*. University of Chicago Press, Chicago.
- Fuentes, M., M. Kuperman, and V. Kenkre. 2003. Nonlocal interaction effects on pattern formation in population dynamics. *Physical Review Letters* 91:158104.
- Gavrilets, S. 1999a. A dynamical theory of speciation on holey adaptive landscapes. *American Naturalist* 154:1–22.
- . 1999b. Dynamics of clade diversification on the morphological hypercube. *Proceedings of the Royal Society of London B* 266:817–824.
- Gould, S. 1989. *Wonderful life: the burgess shale and the nature of history*. Norton, New York.
- Hansen, T., and E. Martins. 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* 50:1404–1417.
- Harmon, L., J. Schulte II, A. Larson, and J. Losos. 2003. Tempo and mode of evolutionary radiation in iguanian lizards. *Science* 301:961–964.
- Harris, T. 1963. *The theory of branching processes*. Springer, Berlin.
- Harvey, P., R. May, and S. Nee. 1994. Phylogenies without fossils. *Evolution* 48:523–529.
- Houchmandzadeh, B. 2002. Clustering of diffusing organisms. *Physical Review E* 66:052902.
- Hutchinson, G. E. 1959. Homage to Santa Rosalia, or why are there so many kinds of animals? *American Naturalist* 93:254–259.
- Jagers, P. 1975. *Branching processes with biological applications*. Wiley, London.
- Jernvall, J., J. Hunter, and M. Fortelius. 1996. Molar tooth diversity, disparity, and ecology in cenozoic ungulate radiations. *Science* 275:1489–1492.
- Kauffman, S. 2000. *Investigations*. Oxford University Press, Oxford.
- Kendall, D. 1948. On the generalized “birth and death” process. *Annals of Mathematical Statistics* 19:1–15.
- Kimura, M. 1983. *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge.
- Le Gall, J. 1999. *Spatial branching processes, random snakes, and partial differential equations*. Birkhauser, Basel.
- Lynch, M. 1990. The rate of morphological evolution in mammals from the standpoint of neutral expectation. *American Naturalist* 136:727–741.
- Martins, E. 1994. Estimating the rate of phenotypic evolution from comparative data. *American Naturalist* 144:192–209.
- Nee, S., R. May, and P. Harvey. 1994. The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London B* 344:305–311.
- Orr, H. 2000. Adaptation and the loss of complexity. *Evolution* 54:13–20.
- Plotkin, J., J. Chave, and P. Ashton. 2002. Cluster analysis of spatial patterns in Malaysian tree species. *American Naturalist* 160:629–644.
- Raup, D. 1966. Geometric analysis of shell coiling: general problems. *Journal of Paleontology* 40:1178–1190.
- . 1967. Geometric analysis of shell coiling: coiling in ammonoids. *Journal of Paleontology* 41:43–65.
- Raup, D., and S. J. Gould. 1974. Stochastic simulation and evolution of morphology: towards a nomothetic paleontology. *Systematic Zoology* 23:305–322.
- Rueber, L., E. Verheyen, and A. Meyer. 1999. Replicated evolution of trophic specializations in an endemic cichlid fish lineage from lake tanganyika. *Proceedings of the National Academy of Sciences of the USA* 96:10230–10235.
- Saunders, W., and A. Swan. 1984. Morphology and morphologic diversity of mid-Carboniferous (Namurian) ammonoids in time and space. *Paleobiology* 10:195–228.
- Sepkoski, J., and A. Miller. 1998. Analysing diversification through time. *Trends in Ecology & Evolution* 13:158–159.
- Shnerb, N. 2004. Pattern formation and nonlocal logistic growth. *Physical Review E* 90:061917.
- Skellam, J. 1951. Random dispersal in theoretical populations. *Biometrika* 38:196–218.
- Slatkin, M. 1981. A diffusion model of species selection. *Paleobiology* 7:421–425.
- Stauffer, D., and A. Aharony. 1992. *Introduction to percolation theory*. 2nd ed. Taylor & Francis, London.
- Swan, A., and W. Saunders. 1987. Function and shape in Late Paleozoic (mid-Carboniferous) ammonoids. *Paleobiology* 13:297–311.
- Thomas, R., R. Shearman, and C. Stewart. 2000. Evolutionary exploitation of design options by the first animals with hard skeletons. *Science* 288:1239–1242.
- Valentine, J. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. *Palaeontology* 12:684–709.
- . 1980. Determinants of diversity in higher taxonomic categories. *Paleobiology* 6:444–450.
- . 1986. Fossil record of the origin of Bauplaene and its implications. Pages 209–231 in D. Raup and D. Jablonski, eds. *Patterns and processes in the history of life*. Springer, Berlin.
- . 1995. Why no new phyla after the Cambrian: genome and ecospace hypotheses revisited. *Palaios* 10:190–194.
- Valentine, J., and D. Erwin. 1987. Interpreting great developmental experiments: the fossil record. Pages 71–107 in R. Raff and E. Raff, eds. *Development as an evolutionary process*. Liss, New York.
- Valentine, J., and T. Walker. 1986. Diversity trends within a model taxonomic hierarchy. *Physica D* 22:31–42.
- van den Bosch, F., J. Metz, and O. Diekmann. 1990. The velocity of spatial population expansion. *Journal of Mathematical Biology* 28:529–565.
- van Kampen, N. 2001. *Stochastic processes in physics and chemistry*. Elsevier Science, Amsterdam.
- Wagner, G., and L. Altenberg. 1996. Perspective: complex adaptations and the evolution of evolvability. *Evolution* 50:967–976.
- Walker, T., and J. Valentine. 1984. Equilibrium-models of evolutionary species-diversity and the number of empty niches. *American Naturalist* 124:887–899.
- Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress of Genetics* 1:356–366.
- Young, W., A. Roberts, and G. Stuhne. 2001. Reproductive pair correlations and the clustering of organisms. *Nature* 412:328–331.

Associate Editor: Douglas H. Erwin
Editor: Jonathan B. Losos